DOCUMENT RESUME

ED 419 002 TM 028 299

AUTHOR Stansfield, Charles W.
TITLE Experiences and Issues Related to the Format of Bilingual Tests: Dual Language Test Booklets versus Two Different Test Booklets.
PUB DATE 1997-01-00
NOTE 8p.
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *Bilingual Education; Elementary Secondary Education; Foreign Countries; *Limited English Speaking; *Native Speakers; *Test Construction; Test Format
IDENTIFIERS *Parallel Test Forms

ABSTRACT

This report, requested by the Massachusetts Department of Education, addresses the issue of the format of native language assessments of students' educational achievement. There are two basic options when administering a translation or adaptation of a test in another language: one is to produce test booklets in both languages and then determine which booklet should be used by the examinee; the other option is to produce the booklet in a format that uses parallel columns. A request for information from language testing specialists who subscribe to a listserv found some information about the use of these and other approaches. In Mexico, tests are administered to Mayan students in Maya and Spanish, and the student receives the higher score earned. Bilingual formats are not used in Micronesia; the teacher or administrator selects the language of the test. The Sylvan Prometric company allows students to access both languages when taking a computer administered test. The Human Sciences Research Council of South African uses a format with items printed in tandem in both languages. In Finland, the Swedish speaking minority may choose a Swedish language test by declaring it in advance. A similar approach is used in Rhode Island, where students may choose minority language versions when available. The Chinese Proficiency Test of the Center for Applied Linguistics uses characters used in Taiwan and those used in the People's Republic of China in parallel columns, and examinees may use one or both columns. This brief review suggests that when parallel columns are used, users are satisfied, although some offer well-intentioned and possibly valid reasons to avoid this approach. A concern when separate versions are used is that there may be pressure to use the dominant language. If the parallel column format is chosen, students should be allowed some additional time to take the test. (Contains two references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 419 002

Experiences and Issues Related to the Format of Bilingual Tests: Dual Language Test Booklets versus Two Different Test Booklets

Author:
Charles W. Stansfield

TM028299

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Charles Stansfield

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Experiences and Issues Related to the Format of Bilingual Tests: Dual Language Test Booklets versus Two Different Test Booklets

Charles W. Stansfield
Second Language Testing, Inc.

Background. On November 29, 1996, staff of the Assessment Unit of the Massachusetts Department of Education and staff at Advanced Systems in Measurement and Evaluation requested that Charles Stansfield write a brief report on the issue of the format of native language assessments of students' educational achievement. The motivation for the request was to learn from the experience of others involved in assessing student achievement across languages.

There are two basic options when administering a translation or adaptation of a test in another language. One is to produce test booklets in both languages and then determine which booklet should be used by the examinee. Another is to produce the test booklet in a format that uses parallel columns; for example, with the left column in English and the right column in Spanish. Other variations are also possible, including putting the two languages on facing pages or having students take the test in both languages and giving the student the higher of the two scores.

Related to the issue is whether additional time should be allotted to examinees who take the test in the parallel column format.

In order to investigate these options, a computerized search of the ERIC database was conducted using descriptors relating to testing, bilingualism, language processing, and reading skills. While this produced a considerable number of abstracts, the search did not produce any articles that dealt directly with the subject of this report. Several potentially relevant articles were scanned, but again no direct information on the matter was found.

In addition, a request for input was transmitted to 250 language testing specialists who subscribe to a listserve. Over half of these subscribers live and work outside of the United States. The request for input produced several interesting comments and descriptions of local practice in different parts of the world. This report describes the input received and the information found that is relevant to the problem.

Mexico. Harold Ormsby, a professor of applied linguistics at the National Autonomous University (UNAM) in Mexico City, responded over the internet with a description of the general achievement testing program he is involved in in the Yucatan Peninsula. These tests, which are under development, are in Maya and Spanish. Ormsby and his colleagues at the UNAM and at the Indian Education Subdirectorate of the Ministry of Education of the Yucatan, have

decided that students should take the test in both languages (Maya and Spanish). They believe that if allowed to choose the language of the test, students will be influenced by expectations that they should take the test in Spanish. Students may sense peer pressure against taking the test in another language. Related to this may be the parents' expectation that the students take the test in the societally dominant language (Spanish). Many parents, although essentially monolingual Maya speakers themselves, would like to see their children progress in the societally dominant language.

Ormsby and his colleagues believe that the students should receive the higher of the two scores. The assumption here is the higher score is a more accurate representation of the student's true ability level.

Timing. In Mexico, parents and children's advocates react bitterly to time limits on tests. With the Spanish and Maya tests, a basic time limit is set based on the estimated amount of time it will take 85% of the examinees to finish the test. This amount of time will be announced at the beginning of the test, and students will be told that they are expected to finish in that amount of time. Then during the test, students will be actually allowed an additional 10 minutes. By the end of the time period, 95% of the students should have turned in their answer sheets. Those who have not completed the test will be dealt with individually. This individual treatment may involve the teacher scanning the answer sheet and talking with the student to see if the student is truly progressing. If the teacher thinks the student can profit from more time, then the teacher can decide how much more time the student will be given, when, and where. The tests will be administered separately in each language on different days.

Ormsby and his colleagues considered and rejected the parallel column approach because of fear that students would react differently when faced with it. They felt that some would read each item word-by-word in both languages several times. Others would efficiently choose one language and stick with it. Those that choose to read each item twice would take far longer to complete the test, with the result that their scores might not be comparable with efficient test takers, unless they were given unlimited time to finish. The fear was that such unlimited time might prove impractical for test administrators.

Micronesia. Bilingual formats are not used in Micronesia. If the test is available in more than one language, the teacher or administrator selects the language of the test. Dan Robinson, director of English Language Institute and Mary Spencer, director of the Micronesian Languages Institute (which develops proficiency and achievement tests in Micronesian languages) at the University of Guam feel that the facing-page approach essentially measures the students' ability to use knowledge and clues from both languages to get the answer. Thus, their score might be different that it would

be in only one language. They feel that this is not the way people normally operate. That is, usually, one has to operate in one language or another.

Sylvan Prometric. Sylvan Prometric is the world's largest provider of computer administered tests. It administers tests for multiple clients. Each client has the option of offering the tests in other languages. This is particularly true for its technology certification tests that are offered throughout the world. The examinees are highly educated adults. The tests are typically 90 minutes in length, but when they are administered in both languages, the time limit is increased by 30 minutes. When both languages are an option, the student can tile (display) the two screens and view them simultaneously, or he can cascade them (impose one over another), and even size them on the screen. Sylvan has no notion of whether examinees view this as advantageous, but they assume that examinees like being able to access the item in both languages. These are all multiple-choice items.

South Africa. The Human Sciences Research Council (HSRC), a semi-autonomous research institution, has traditionally been the prime test development institution in South Africa. The country now has eleven official languages, nine autoctonous languages plus Africaans and English. Typically, the HSRC constructs the test in a bilingual format, with the two languages used being Africaans and English. More recently, an indigenous language, such as Zulu, is paired with English.

The HSRC divides the page into two columns with the items printed in tandem in both languages. Examinees can easily see both languages and this is viewed positively by the testing agency. No complaints about the format have been received from students or teachers. The informal impression of one HSRC researcher (Jeff Chamberlain) is that the format is viewed as fair in that it gives the candidate additional information if the question is not understood clearly in one of the languages. Tests are pretested in this format and timing is set as the amount of time necessary for 90% to finish the test. Examinees are not asked which column they have used, or whether they use both. However, at the beginning of the test there is a question asking the examinee to mark on the answer sheet the language in which he or she prefers to be tested. This mark is used to divide the answer sheets into two groups, with each group analyzed separately using traditional item and test statistics.

Finland. Finland has two official languages: Finnish and Swedish. Finland used to belong to Sweden, and it still has a small native-Swedish-speaking minority. Nearly all the Swedish speaking population speaks Finnish very well, especially in Helsinki, but few Fins speak Swedish well, even though its study is compulsory in school.

BEST COPY AVAILABLE

Finland has administered high school exit exams for about 100 years. Some high school seniors are bilingual and biliterate. The exit exams are available in both Finnish and Swedish. Test tasks in math, science, humanities, and languages are all the same in each language, or as close to the same as possible. Usually, the tests are administered in one language, according to the students mother tongue and the language in which a student has been educated. So, if a student attends a Swedish school, the tests are normally in Swedish.

One basic principle in Finland concerning the language of the test is that the student must declare in advance in which language he or she wants to take the test. The student may not change languages after the test has begun.

Rhode Island. The Rhode Island Department of Education (RIDE) began administering tests that are part of its state assessment system in four non-English languages in the Spring of 1996. The languages are Spanish, Portuguese, Lao, and Khmer (referred to as Laotian and Cambodian in Rhode Island). The students were given one test booklet in each language and were forced to write their answers in one test booklet. Thus, students had to determine the language in which they wanted to take the test at the start, or very soon afterwards. No negative feedback was reported to the state on this procedure. Far fewer students than expected chose to respond in the non-English language. This may or may not be related to the issue of test format.

The Chinese Proficiency Test. The Center for Applied Linguistics (CAL) has developed a Chinese Proficiency Test (CPT). The CPT is administered to students of Chinese language throughout the US. The testing program began in 1986, and roughly 1,000 students per year take the test.

The CPT is a test of listening, reading, and grammar. The reading and grammar sections employ the parallel column format. Chinese characters are represented in two ways on the same page. Traditional characters (used in Taiwan) are in the left column and simplified characters (used in the Peoples Republic of China) are in the right column. Ethnic Chinese who take the test, as well as students who are not ethnic Chinese, may have been exposed to one writing system, the other, or both. As a result, students taking the test appreciate the availability of the parallel column format. In practice, depending on their familiarity with one or both character systems, an examinee may use one or both columns to determine the meaning of a given CPT item stem or option.

CAL did not intentionally set the time limits for the test to allow for the reading of both columns; yet undoubtedly, a number of examinees use both columns to ascertain meaning. Chinese characters are somewhat ideographic. Thus, if a student does not get the meaning from one character, he or she may look at the

character in the parallel column to see if they recognize it, or to see if it gives them a clue to the meaning. Examinees are able to do this rapidly because they are working within the same language and with parallel writing systems. About 50% of the characters are identical across the two writing systems. The fact that examinees are able to do this successfully in Chinese does not mean that no additional time would be required to process information in two different languages. English and Spanish for example are quite different in the way they show the organization of elements within a sentence. English relies largely on word order while Spanish uses a more flexible word order but relies largely on word endings.

Other related issues.

Use of a bilingual dictionary. Studies of dictionary use during an English examination (Bousoussan 1983; Bensoussan, Sim, & Weiss 1985) found that bilingual dictionaries do not significantly improve test scores. However, they do add the amount of time required to complete the test. This finding would seem to support the proposition that the bilingual (parallel column) format would require additional test administration time, since it also would encourage the examinee to do some additional reading during the test. Similarly, the administrative accommodation of allowing use of a bilingual dictionary would also seem to warrant the administrative accommodation of some additional testing time.

Conclusion. This brief review of the practices and experiences of others leads to the following very tentative conclusions. 1.) When the parallel column or a similar format has been tried, people seem to be satisfied with it. 2.) When the parallel column format is not used, people cite well-intentioned and possibly valid reasons why it shouldn't be. These reasons relate to the authenticity of the test format, fear of differing effects on examinees (i.e., different response sets), etc. 3. The concern about the use of separate test booklets in each language is that pressure from peers, teachers, or society to take the test in the societally dominant language will prevail. 4. If the parallel column format is to be used, some additional time should be allowed in order for examinees to complete the tests.

Delivered: January 27, 1997

References

Bensoussan, M. (1983). Dictionaries and test of EFL reading comprehension. *English Language Teaching Journal*, 37(4), 341-345.

Bensoussan, M., Sim, D., & Weiss, R. (1985). The effect of dictionary usage on EFL test performance compared with student and teacher attitudes and expectations. *Reading in a Foreign Language*, 2(2), 262-276.